Problem 1. Channels with memory have higher capacity. Consider a binary symmetric channel with $Y_i = X_i \oplus Z_i$, where $\oplus$ is mod 2 addition, and $X_i, Y_i \in \{0, 1\}$.

Suppose that $\{Z_i\}$ has constant marginal probabilities $\Pr\{Z_i = 1\} = p = 1 - \Pr\{Z_i = 0\}$, but that $Z_1, Z_2, \ldots, Z_n$ are not necessarily independent. Assume that $(Z_1, \ldots, Z_n)$ is independent of the input $(X_1, \ldots, X_n)$. Let $C = \log 2 - H(p, 1 - p)$. Show that

$$\max_{P_{X_1, X_2, \ldots, X_n}} I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n) \geq nC.$$ 

Problem 2. Consider two discrete memoryless channels. The input alphabet, output alphabet, transition probabilities and capacity of the $k$th channel is given by $X_k$, $Y_k$, $P_k$ and $C_k$ respectively ($k = 1, 2$). The channels operate independently. A communication system has access to both channels, that is, the effective channel between the transmitter and receiver is $X_1 \times X_2$, output alphabet $Y_1 \times Y_2$ and transition probabilities $P_1(y_1|x_1)P_2(y_2|x_2)$. Find the capacity of this channel.

Problem 3. Let $P_1$ and $P_2$ be two channels of input alphabet $X_1$ and $X_2$ and of output alphabet $Y_1$ and $Y_2$ respectively. Consider a communication scheme where the transmitter chooses the channel ($P_1$ or $P_2$) to be used and where the receiver knows which channel was used. This scheme can be formalized by the channel $P$ of input alphabet $X = (X_1 \times \{1\}) \cup (X_2 \times \{2\})$ and of output alphabet $Y = (Y_1 \times \{1\}) \cup (Y_2 \times \{2\})$, which is defined as follows:

$$P(y, k'|x, k) = \begin{cases} P_k(y|x) & \text{if } k' = k, \\ 0 & \text{otherwise.} \end{cases}$$

Let $X = (X_k, K)$ be a random variable in $X$ which will be the input distribution to the channel $P$, and let $Y = (Y_k, K) \in Y$ be the output distribution. Define $X_1$ as being the random variable in $X_1$ obtained by conditioning $X_k$ on $K = 1$. Similarly define $X_2$, $Y_1$ and $Y_2$. Let $\alpha$ be the probability that $K = 1$.

(a) Show that $I(X; Y) = h_2(\alpha) + \alpha I(X_1; Y_1) + (1 - \alpha)I(X_2; Y_2)$.

(b) What is the input distribution $X$ that achieves the capacity of $P$?

(c) Show that the capacity $C$ of $P$ satisfies $2^C = 2^{C_1} + 2^{C_2}$, where $C_1$ and $C_2$ are the capacities of $P_1$ and $P_2$ respectively.

Problem 4. Show that a cascade of $n$ identical binary symmetric channels,

$$X_0 \rightarrow \text{BSC #1} \rightarrow X_1 \rightarrow \cdots \rightarrow X_{n-1} \rightarrow \text{BSC #n} \rightarrow X_n$$

each with raw error probability $p$, is equivalent to a single BSC with error probability $\frac{1}{2}(1 - (1 - 2p)^n)$ and hence that $\lim_{n \to \infty} I(X_0; X_n) = 0$ if $p \neq 0, 1$. Thus, if no processing is allowed at the intermediate terminals, the capacity of the cascade tends to zero.
**Problem 5.** Consider a memoryless channel with transition probability matrix $P_{Y|X}(y|x)$, with $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. For a distribution $Q$ over $\mathcal{X}$, let $I(Q)$ denote the mutual information between the input and the output of the channel when the input distribution is $Q$. Show that for any two distributions $Q$ and $Q'$ over $\mathcal{X}$,

(a) \[ I(Q') \leq \sum_{x \in \mathcal{X}} Q'(x) \sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) \log \left( \frac{P_{Y|X}(y|x)}{\sum_{x' \in \mathcal{X}} P_{Y|X}(y|x') Q(x')} \right) \]

(b) \[ C \leq \max_x \sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) \log \left( \frac{P_{Y|X}(y|x)}{\sum_{x' \in \mathcal{X}} P_{Y|X}(y|x') Q(x')} \right) \]

where $C$ is the capacity of the channel. Notice that this upper bound to the capacity is independent of the maximizing distribution.